

Technological Impacts on Return to Education in Brazil
Yirui Zhao

Michelle Connolly, Faculty Advisor
Rafael Dix-Carneiro, Faculty Advisor
Daniel Xu, Faculty Advisor

*Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in
Economics in Trinity College of Duke University.*

Duke University
Durham, North Carolina
2023

Table of Contents

Acknowledgements.....	3
Abstract.....	4
Introduction.....	5
Literature Review.....	10
Empirical Framework.....	12
Data.....	16
Results.....	22
Discussions.....	26
Conclusion.....	28
References.....	29
Appendix.....	31

Acknowledgements

I like to thank Professor Michelle Connolly. This thesis cannot be done without her continuous help on my research methods and revisions.

I am also grateful to Professor Rafael Dix-Carneiro for providing me access to the key datasets I needed for my research and guidance of how to use them. The resources provided by him made my research goal achievable.

I also like to thank Professor Daniel Xu. His suggestions and comments helped me solve some key issues I faced when constructing my empirical framework.

Finally, I would like to thank all my peers in the honor seminar. The discussions with them were inspiring and their feedbacks made my research more comprehensive.

I take full responsibility for the thesis including any possible errors.

Abstract

The wage return to education has been studied for a long time. Acemoglu and Autor (2010) connect the decrease of medium-level job opportunities in the U.S. with technological advances. Their theoretical model predicts that if technology replaces routine jobs, workers with medium-level skills will experience decreases in wages relative to both high-skill workers (who become more productive with the improved technology) and low-skill workers (who can less easily be replaced since their work is not routine). Moreover, their theoretical model predicts that if medium-skill workers are closer substitutes for low-skill workers than they are for high-skill workers, the relative return of high-skill workers to low-skill workers should increase. Using education as proxy of skill (Acemoglu & Autor, 2012), this paper checks if these three predictions about relative wage returns to education also hold in Brazil. This paper finds that the impact of technological change on the Brazilian formal labor market between 1986 and 2010 is consistent with predicted changes in the return to education for medium-skill workers relative to both low and high skill workers. The impact is consistent with predicted changes in the return to education for high-skill workers relative to low-skill workers when Lula's presidency is considered in the model.

JEL classification: J24; J31; O33

Keywords: education; skill premium; income inequality; occupation; technology

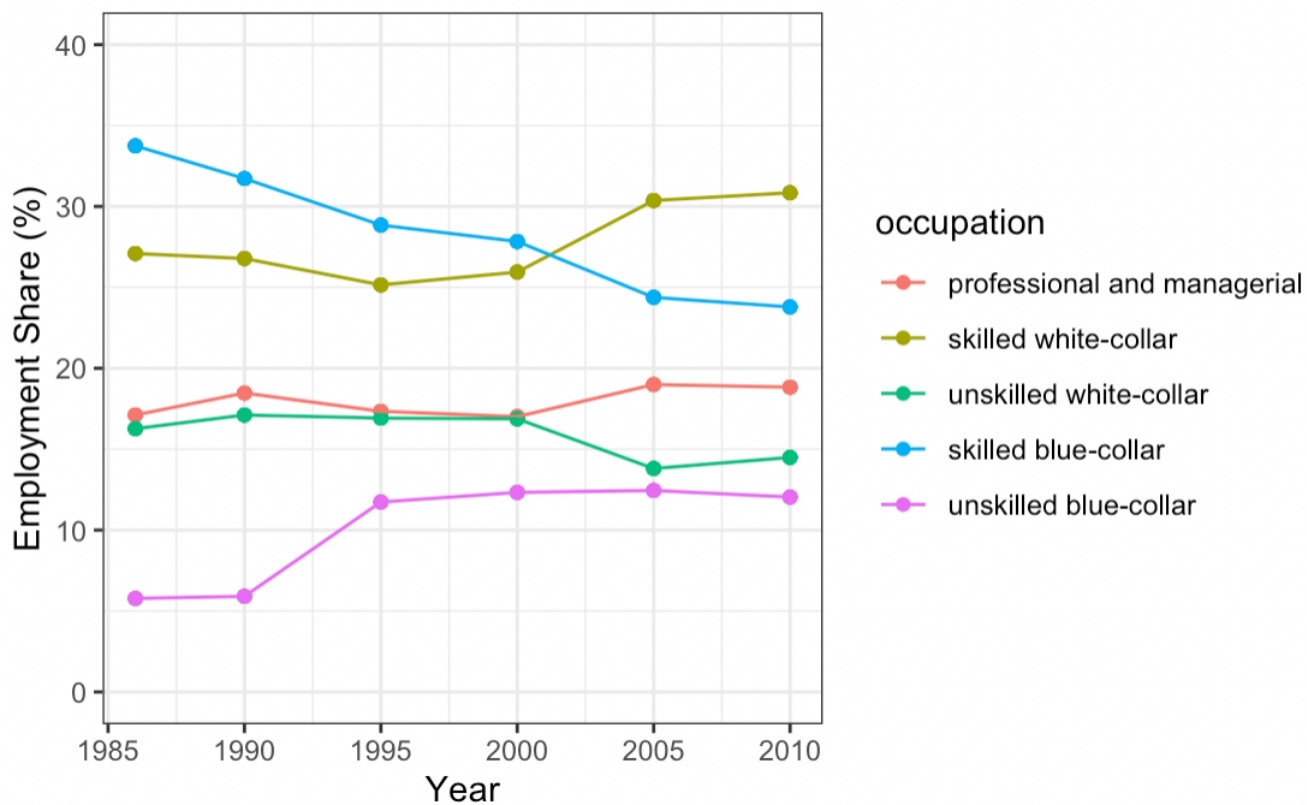
I. Introduction

The wage return to education has been discussed for a long time since Mincer constructed a function to explain income with schooling and experiences. Just like the prices of goods, the return to education is determined by demand and supply. In the US, there has been both an increase in relative supply of college educated workers, and an increase in relative wage return of a college degree to all other workers (Acemoglu & Autor, 2010). This suggests something in the labor market is pushing the demand of highly educated workers upward. Studies have argued that skill-biased technology is the factor that is shaping the labor market demand and bringing skill premiums to workers with high education levels (Acemoglu & Autor, 2010; Berman et al., 1998). Moreover, observations from the US suggest that technology is replacing routine jobs generally filled by people with medium education levels. This has caused job polarization where the opportunities for mid-wage workers are decreasing, while opportunities for workers at the two ends of the job market (low-wage and high-wage) are increasing (Keller & Utar, 2016). Decreased demand for these medium-skill-level jobs further pushes people with medium education levels – who are already working to shift to jobs with lower education requirements (Acemoglu & Autor, 2010). All these trends suggest that technology changes in the US are differentially impacting the return to different education levels. According to Acemoglu and Autor (2010), the technology improvements started in the 1970s is primarily harming people with medium-level educations.

Data from the Brazilian RAIS dataset, which is an individual-level panel data documenting detailed information about workers in the Brazilian formal labor market, further indicates that this employment trend is not unique to developed countries. According to Figure 1, the employment shares of unskilled white-collar and skilled blue-collar jobs in the Brazilian formal labor market shrank from 1986 to 2010, especially after 1996, while the employment shares of professional and managerial,

skilled white-collar, and unskilled blue-collar jobs increased. I consider professional and managerial, and skilled white-collar jobs as high-skill jobs, unskilled white-collar and skilled blue-collar jobs as medium-skill jobs, and unskilled blue-collar jobs as low-skill jobs. This observation shows that jobs that require medium level of skills are losing their shares in the Brazilian labor market, which is similar to what people observed in the US.

Figure 1: Employment shares by major occupation groups, Brazil 1986-2010

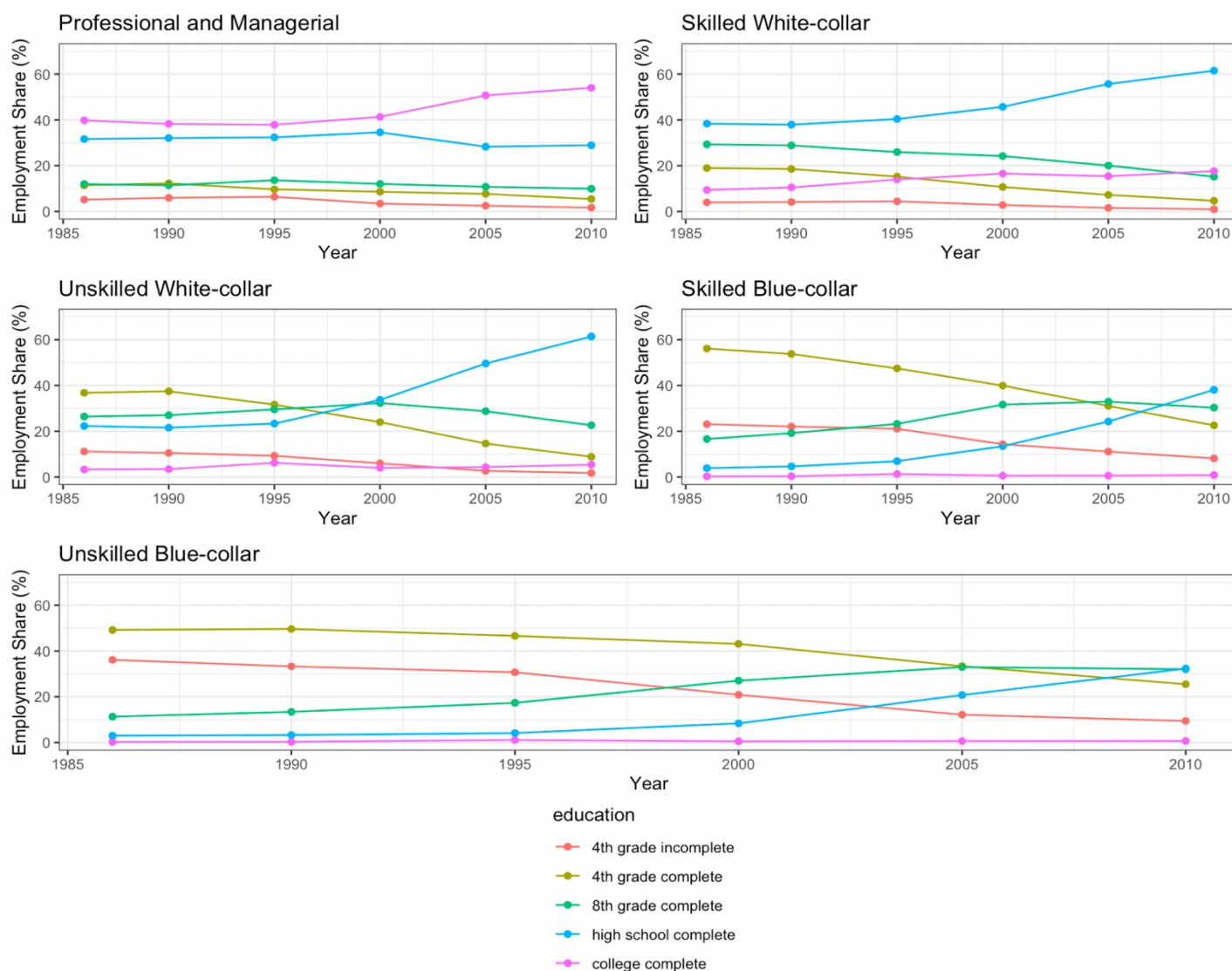


Source: RAIS for years 1986, 1990, 1995, 2000, 2005, and 2010.

According to Figures 2, from 1986 to 2010, all occupations in the formal labor market have increase in their share of workers with higher education credentials. This can be explained by two factors. First, as will be shown later in section IV, the education levels of workers in the formal labor

market generally increases across time. Still, this does not fully explain what Figures 2 is showing. For example, the relative amount of workers who completed 8th grade decreased from 2000 to 2010. However, within unskilled blue-collar occupations, workers who completed 8th grade had a steady increase in proportion during that time period, which suggests they are shifting to jobs which require lower skills.

Figure 2: Occupational employment shares by worker education, Brazil 1986-2010



Source: RAIS for years 1986, 1990, 1995, 2000, 2005, and 2010.

This tendency demonstrates that when original jobs are no longer available, workers with a particular education level are closer substitutes for workers employed in occupations that require lower skills (education levels) than they are for workers employed in occupations that require higher skills (education levels). That is, in Brazil's case, when the employment shares of unskilled white-collar and skilled blue-collar jobs are shrinking, workers that originally worked or would have worked in these occupations are more likely to shift to unskilled blue-collar jobs, rather than professional and managerial jobs.

With the tendencies mentioned above, the employment share changes, and study from Acemoglu and Autor (2010), it is reasonable to predict that in Brazil the relative wage return of medium-level educations should decrease relative to both low and high education levels and the relative wage return of high-level educations to low-level educations should increase because of technological improvements.

This paper first demonstrates that technological improvements (proxied by high tech imports as a share of GDP) in Brazil are indeed harming workers with medium-level educations the most in relative terms. Given this premise, this paper then checks if the wage inequalities between high and low education workers shifts in the same directions as Acemoglu and Autor predicted. The results are consistent with their predictions about changes in relative wage between low-skill and medium-skill workers, and changes in relative wage between medium-skill and high-skill workers. However, the results suggest the wage gap between low-skill and high-skill workers in Brazil decreases with technology improvements. This may be biased by Lula's presidency during 2003-2010 since Lula conducted many economic reforms to reduce overall inequality at that period. After I add a dummy variable indicating Lula's presidency to regressions, the results suggest the wage gap between low-skill and high-skill workers in Brazil increases with technology improvements. All other findings remain the

same. In the end, the paper discusses the possibility that job polarization can be transferred from developed countries to developing countries as technological advances are adopted in Brazil.

Section II of this paper goes through the previous studies of returns to education and a more detailed explanation of Acemoglu and Autor's theory. Section III presents the construction of my empirical model based on Mincer's function. Section IV describes my data. Section V and VI presents and discusses my findings. Section VII concludes.

II. Literature Review

A. Studies on Returns to Education

Returns to education have different meanings in microeconomic and macroeconomic contexts. From the macro perspective, returns to education represents the impact of education on national economic growth (Sianesi & Reenen, 2003). From the micro perspective, education is a private investment in “human capital” and the return is reflected through individual earnings (Harmon et al., 2003). The return to education in this paper is consistent with the definition of the microeconomic perspective.

When education is viewed as an investment, its return is related with its risk. However, it is noticed that the return of education is higher than other investments with similar risks (Harmon et al., 2003). This implies there exist barriers which keep people from receiving their optimal amount of education, such as from high school to college (Harmon et al., 2003). Moreover, school quality can impact returns to the same education level, while technology changes can impact returns to different education levels (Berman et al., 1998; Card & Krueger, 1992; Goldin & Katz, 2010).

Perceived returns to education also affect people’s schooling decisions (Harmon et al., 2003; Jensen, 2010; Sianesi & Reenen, 2003). That is, people decide whether they continue their study or not based on the amount of return they think they will get after graduation. While their estimated return can be wrong, students with a higher estimated return from schooling generally intend to continue their study for a longer time (Jensen, 2010).

B. The Canonical Model and Improvements

The traditional Canonical Model on technology and returns to education assumes that there are two skill groups, low-skill and high-skill. Technology is assumed to be complementary to either high-skill workers or low-skill workers. The model predicts that when technology is “skill-biased” and resulting demand increases faster than the supply of high-skill workers, inequality between low-skill and high-skill workers rises, and vice versa if supply increases faster than demand (Acemoglu & Autor, 2012).

Acemoglu and Autor extend this model by adding medium-skill workers and follow the traditional model assumption that technology can both be used to complement workers and replace workers depending on type (Acemoglu & Autor, 2010). They suggest that recent technological change has been able to replace routine tasks: tasks previously fulfilled by medium-skilled workers. For example, travel agents are replaced by online reservations. But that change does not replace non-routine tasks previously fulfilled by low-skill and high-skill workers, such as cleaning and managerial positions respectively. Hence the relative wage of medium-skilled workers to low-skill workers decreases and the relative wage of high-skill workers to medium-skill workers increases. Moreover, if medium-skill workers are closer substitutes to low-skill workers than high-skill workers, the wage inequality between high-skill workers and low-skill workers rises, all else equal (Acemoglu & Autor, 2010). Their model is driven by observations of the US labor market. Observing if this pattern also holds in a developing country is a good way to see how well their model can be applied to different labor markets.

III. Empirical Framework

The common way to estimate the return of education is using the Mincer's human capital earnings function (Mincer, 1974):

$$\log w_i = \alpha + \mathbf{X}_i\beta + rS_i + \delta E_i + \gamma E_i^2 + u_i \quad (1)$$

Where w is an earning measure for individual i . S is a measure of education, that is, the number of years an individual spent in school. E is work experience in years. The equation also contains a quadratic form for the work experience measure in order to reflect the concavity of returns to experience. Mincer also suggests a proxy for work experience when direct information on experience is unavailable. Assuming an individual of age A started school at age T , finished S years of schooling in exactly S years (no early or late graduation), and began working immediately after the graduation from school, then the years he or she spent working can be expressed as $E_i = A - S - T$. X is a set of other variables besides education and job experience that is assumed to affect earnings and u is the random disturbance term assumed to be independent of X and S .

Mincer's original education variable is measured in years. However, the RAIS dataset used in this paper documents education in credentials, such as high school complete and college complete. The measure in credentials is better than the measure in years as the return of schooling is not linear. Fulfilling a particular year of schooling, such as the last year in high school or college can bring extra wage premium (Hungerford & Solon, 1987; Park, 1996). As a result, the way education is documented in the RAIS dataset can be an advantage. Hence, here the education variable documents an individual's highest credential, which can be 4th grade incomplete, 4th grade complete, high school complete, and college complete. Workers who completed 8th grade (medium-level education) are used as the reference

group in order to easily track the change in their wage returns relative to higher and lower education groups across time. The RAIS dataset contains data from 1986 to 2010. Since children began school at age 7 before 2010, work experience can be calculated as $Ei = A - S - 7$ (Stanek 2013). It is assumed that for workers who completed 4th grade, $S = 4$. Similarly, $S = 8$ for workers who completed 8th grade, $S = 11$ for workers who completed high school (high school in Brazil takes 3 years), and $S = 16$ for workers who completed college (as Bachelor programs in Brazil are usually 4-6 years).

Besides education and work experience, previous studies have shown many other factors that are related with earning outcomes which can be considered in terms of variables in X – the set of additional traits related to earnings. Wage gaps by gender are well documented and have had much research done attempting to explain the reasons for this gap. Regional income disparities are observed over the period 1985-2001 in Brazil (Silveira-Neto & Azzoni, 2006). As a result, I include gender and region as variables in X . I would also like to include race and nationality in X because previous studies suggest the lack of qualifications and incomplete assimilation generate a wage gap between immigrants and native workers (Nielsen et al., 2004). Racial income differences are also observed. However, the RAIS dataset does not include a race variable until 2004, and most of the workers in the formal labor market are Brazilians. So, the regression here does not include race in order to see the changes of return to education through a longer time period and does not include nationality given the data is very unbalanced in this variable.

Since the study aims to show the impacts of technology improvements to workers with different education levels, technology is also an important component of earnings outcomes. We can view the impacts of technology in two ways: the first is the general impact of technology improvements to all the workers in the formal labor market, and the second is specific impacts to workers with different education credentials. As a result, I include one variable proxying the general impact of technology and

further interact this term with different education levels to see its impacts on different education groups. I use values of high-tech imports to Brazil as a share of total GDP each year as a proxy for the inflow/impact of technology from abroad, which is inspired by Connolly (2003)¹. Also, since the job market needs time to respond to the technology change, I use a one-year lag when studying relationship between wage and technology change.

Brazil's economy was not very stable during my period of study. I therefore also include Brazil's annual real GDP per capita in my empirical model to capture the influence of general macro-economic conditions in Brazil.

Taking my research goal, previous studies, and available data together, the final empirical model is shown below.

$$\begin{aligned} \log w_{it} = & \alpha + \beta_1 T_{t-1} + \sum_k \beta_{2k} S_{kit} + \sum_k \beta_{3k} (S_{kit} \times T_{t-1}) \\ & + \beta_4 E_{it} + \beta_5 E_{it}^2 + \beta_6 G_{it} + \beta_7 GCAP_t + \sum_l \beta_{8l} L_{lit} + u_{it} \end{aligned} \quad (2)$$

W is real December monthly wage in year t for worker i in year t . I use log of real wage to reduce skewness of it. S is the education measure as k level dummies, including 4th grade incomplete, 4th grade complete, 8th grade complete, high-school complete, and college complete, with workers who completed 8th grade as the reference group. E is work experience in years. Technology is represented by the value of high-tech imports as a share of total GDP in each year. G is the gender dummy variable which equals to 1 for female workers. $GCAP$ is the annual real GDP per capita. L is the location measure as l level dummies, including Northeast, Southeast, South, and Center West. Workers who work in plants in the

¹ Since most innovation occurs in developed countries, technology must diffuse to developing countries. This can occur through multiple channels, including multinational, worker migration, international trade in capital goods embodying tech, etc. Here I just consider one channel because of available data, which is high-tech imports to Brazil as a share of total GDP each year.

Center West belong to the reference group. Again, the benchmark for this regression are male workers in Center West who have completed the 8th grade.

$$\begin{aligned} \log w_{it} = & \alpha + \beta_1 T_{t-1} + \sum_k \beta_{2k} S_{kit} + \sum_k \beta_{3k} (S_{kit} \times T_{t-1}) \\ & + \beta_4 E_{it} + \beta_5 E_{it}^2 + \beta_6 G_{it} + \beta_7 GCAP_t + \sum_l \beta_{8l} L_{lit} + \beta_9 \text{time index} + f_i + u_{it} \end{aligned} \quad (3)$$

$$\begin{aligned} \log w_{it} = & \alpha + \beta_1 T_{t-1} + \sum_k \beta_{2k} S_{kit} + \sum_k \beta_{3k} (S_{kit} \times T_{t-1}) \\ & + \beta_4 E_{it} + \beta_5 E_{it}^2 + \beta_6 G_{it} + \beta_7 GCAP_t + \sum_l \beta_{8l} L_{lit} + \beta_9 \text{time index} + u_{it} \end{aligned} \quad (4)$$

I will run both fixed-effect regression (equation (3)) and random-effect regression (equation (4)), which are constructed based on empirical model of equation (2). Random-effect regression is the most powerful regression because it will not drop time-invariant variables that I am interested in. But coefficients from it can be biased if there exist latent individual fixed effects that are not included in my regression. A fixed-effect regression can control for latent individual effects so the coefficients from it can be less biased. However, it cannot estimate coefficients on any time-invariant variables, such as a worker's education level. So, running both is an optimal way to capture coefficients I want and detect possible bias in estimations. To the extent that there may be a time trend, like a general increase in technology over time, I add a time index to each regression.

IV. Data

The dataset that is used for studying wage returns to education is the RAIS dataset from the Brazilian Ministry of Labor. To proxy for technological change, I include trade information from the United Nations Comtrade Database and GDP information from the CEPII Gravity Database. I use real GDP per capita data from the World Bank to capture the general macro-economic conditions in Brazil.

A. Education and Formal Labor Market

The study uses the RAIS dataset as the primary data source given it allows the study to track individual workers across time by using the unique PIS IDs in the dataset. The RAIS dataset contains detailed information about 103 million individual employees in the Brazilian formal labor market within period 1986-2010, excluding interns, domestic workers, and other minor employment categories, along with those without signed work cards, including self-employed. Based on the trends observed in the introduction, the employment share of medium-skill occupations started to shrink around late 1990s and early 2000s.

The dataset documents individual wages in two ways. The first one is wage of the worker in December, which is written as multiple of the minimum wage in December of that year. The second one is average wage of the worker over the year or over the employment spell within the plant, which is also written as multiple of the minimum wage in December of that year. This paper uses the first method to represent wages as this ensures that workers' formal employment status and earnings are measured at the same time point, thus avoiding the potential confounding effects on average yearly earnings which may from, for example, different job starting time in a year.

For data summary, I derive a random sample which contains information of 3% of workers who had wage in December documented in RAIS from 1986 to 2010 at least once. I use sampling data here because first the number of observations of RAIS in each year is 30-67 million, which makes tracking individual workers in the whole formal labor market population across time difficult. Second, some of the workers with a high education credential (college or high school complete) have multiple jobs in the formal labor market, and these jobs usually have very different skill requirements. For example, a person who completed college is CEO of a company, while at the same time he may also do an unskilled blue-collar job because he has time and is interested in it. As a result, including all these observations in the study can blur the results. A way to solve this problem is to only consider the job with the highest payment as the observed job for the study. This process requires a lot of computation power so applying it to the whole dataset is not feasible here. After a few rounds of tests and comparisons, I am confident that a random sample of 3% of workers who had wage in December documented in RAIS from 1986 to 2010 at least once (2.6 million) is big enough to represent the population properly.

I observe inconsistencies sometimes happen in the data for gender, age, and education levels. Since RAIS is a panel dataset, it allows me to detect and correct some of these inconsistencies by comparing values across time and using frequency to determine the correct data. I especially highlight the way I correct education levels here. Since children in Brazil began school at age 7 before 2010 and it takes 16 or 17 years of education in total for an individual in Brazil to get a college degree, an individual's education level can keep changing before age 25 and usually remains the same afterwards. If a worker's education level suddenly switches after age 25, it is highly likely that there is a documentation error. In Acemoglu and Autor's model, skill level is a fixed property for each worker (2010). Hence, I consider each worker's most common education level in the dataset after age 25 as his/her ultimate education level and use it in later regressions.

As a result, I drop workers who never exceed 25 and workers who do not have documented education levels after age 25 in the dataset. This step removes data of 0.47 million unique workers (1 million observations without other missing values) from the sample. In order to avoid negative values for work experience, I drop observations where a worker's age is smaller than number of years required to achieve his/her ultimate education level. For example, if I observe a worker aged 17 in 1986 and he/she got a college degree after age 25, I know this worker pursued higher education after 1986 and I delete this observation. That is, the observations I keep are ones where workers achieved their ultimate education levels and would not have further education improvements. This step does not change the number of unique workers in the sample.

After all the cleaning, there are 2.2 million unique workers (18 million observations without missing values) in the sample. I am confident that the cleaning will not bias my research results because the formal labor market employment distributions from my cleaned sample (Table A3 in appendix) is consistent with the population distributions I show in section I: employment share of medium-level jobs shrinks across time. Hence, the research context remains the same.

Below I provide distribution of education levels and summary statistics of real wage in year 1986, 1990, 1995, 2000, 2005, and 2010 grouped by education levels. These statistics are derived from my cleaned sample.

Table 1: Education Distribution (in percentage)

Education Level	1986	1990	1995	2000	2005	2010
4th Grade Incomplete	13.23	13.03	11.61	8.77	6.92	6.08
4th Grade Complete	33.38	32.24	29.24	24.34	20.31	18.77
8th Grade Complete	20.68	21.28	22.14	22.76	22.68	22.76
High School Complete	22.31	23.13	26.35	32.44	37.99	39.68
College Complete	10.38	10.32	10.65	11.69	12.10	12.71
Total Observations	650,930	762,303	824,107	878,125	1,086,342	1,232,365

Resource: Random Sample from RAIS Dataset

Education Level	1986	1990	1995	2000	2005	2010
4th Grade Incomplete	621.47 (778.24)	397.09 (676.53)	473.62 (733.53)	478.84 (706.81)	484.29 (694.23)	603.50 (747.22)
4th Grade Complete	773.56 (1002.92)	500.46 (728.94)	613.28 (876.15)	605.90 (818.42)	582.91 (771.55)	695.36 (857.00)
8th Grade Complete	919.30 (1332.5)	595.43 (943.10)	722.53 (1113.22)	701.04 (989.86)	664.72 (904.98)	790.07 (1026.54)
High School Complete	1364.39 (1837.47)	910.71 (1308.20)	1030.36 (1700.50)	1027.93 (1557.08)	926.28 (1380.48)	1151.80 (1607.42)
College Complete	2941.13 (3046.53)	2106.86 (2353.08)	2853.66 (3308.20)	2982.74 (3442.13)	2877.22 (3660.42)	3497.81 (4229.10)
Total Observations	650,930	762,303	824,107	878,125	1,086,342	1,232,365

Resource: Random Sample from RAIS Dataset

Note: numbers inside (.) are standard deviations and numbers above are means.

Table 1 shows that relatively more and more workers in the formal labor market complete high school and college across time. The median of education levels gradually shifted from 8th grade to high school around early 2000s. For consistency, I consider 8th grade complete as the medium education level for later regressions based on data in Table 1. According to Table 2, the average real wage of workers who completed 8th grade experienced a decrease between 1995 and 2005, which corresponds to the period when employment share of medium-skill occupations experienced a shrink. Workers who completed high school and 4th grade, which are not the two ends of the education credentials, also experienced a real wage decrease in that period. All education levels besides 4th grade incomplete experienced a decrease in average real wage in 2005 and a recover in 2010. These may be explained by the economic downturn happened in early 2000s and later economic reforms which stabilized the economy. The number of observations increases across time because the RAIS dataset has more observations over time.

B. Change of High-tech Imports and GDP

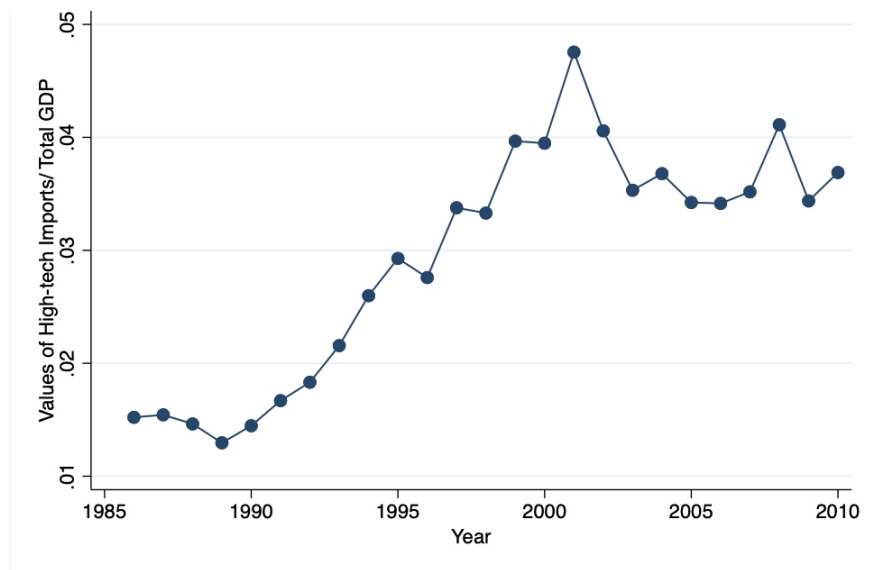
The UN Comtrade Database contains annual and monthly trade statistics by trading partners and commodity types. It covers about 99% trade activities across about 200 countries. In this study, I use annual values of imports reported as Commodity Codes 7, 86, and 89 in the Standard International Trade Classes (SITC) to represent annual values of high-tech imports, as these three codes stand for machinery and transport equipment, scientific and control instruments, photographic goods, clocks, and miscellaneous manufactured goods (such as office equipment, including computers in later years), which are most likely to contain goods involving technology and used by high-tech industries (Connolly, 2003).

I get Brazil's annual total GDP from 1986-2010 from the CEPII Gravity Database to weight the values of high-tech imports as share of total GDP to get rid of inflation problems. The information about annual real GDP per capita in Brazil is from Word Bank.

Figure 3 shows the change of values of high-tech imports as share of total GDP in Brazil from 1986 to 2010. The values started to have a significant increase during the middle of trade liberalization in Brazil (around 1990s). The values experienced a decrease during 2001-2003 and later stabilized. The decrease may be related to the economic downturn in Brazil during that time because of high inflation and currency issues.

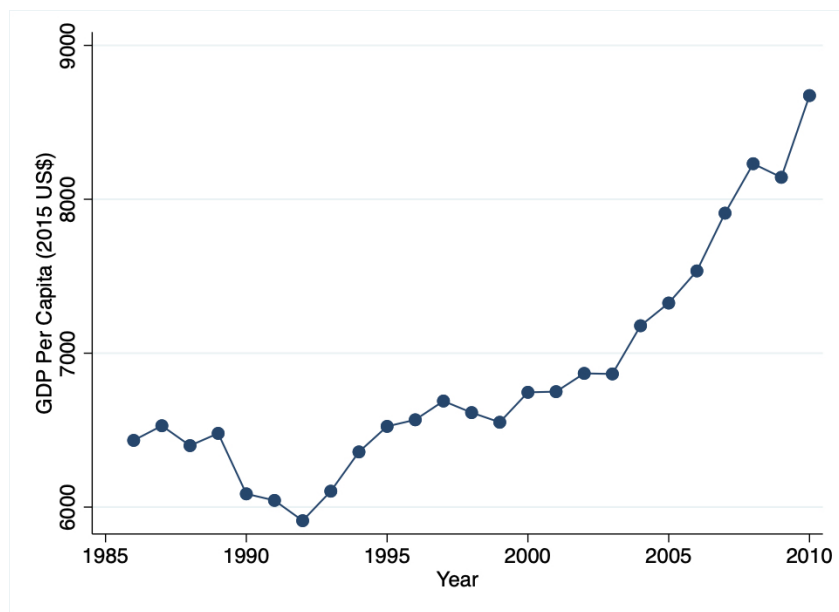
Figure 4 shows the change of real GDP per capita in Brazil from 1986 to 2010. It experienced a drop around late 1980s. This is possibly because of the unstable economy and high inflation during that period. The real GDP per capita started to increase rapidly from 2003, the year when Lula became the president of Brazil. His presidency involved many economic reforms to stabilize the economy and reduce inflation, and that may explain the rapid increase during his ruling period.

Figure 3: Values of High-tech Imports as Share of Total GDP 1986-2010



Source: United Nations Comtrade Database and CEPII Gravity Database

Figure 4: Real GDP Per Capita of Brazil 1986-2010



Source: World Bank

V. Results

Table 3: General Regressions with Fixed Effects, Random Effects, and Year Trend Controls, 1986-2010

Dependent Variable: Log December Real Wage (2005 R\$)	(1) Fixed Effects	(2) Random Effects
Constant	4.399*** (0.004)	3.997*** (0.004)
High Tech Imports/GDP	9.902*** (0.060)	9.954*** (0.058)
4th Grade Incomplete	-	-0.678*** (0.003)
4th Grade Complete	-	-0.320*** (0.002)
High School Complete	-	0.375*** (0.003)
College Complete	-	1.163*** (0.004)
4th Grade Incomplete $\times$ High Tech Imports/GDP	6.305*** (0.102)	6.615*** (0.093)
4th Grade Complete $\times$ High Tech Imports/GDP	1.663*** (0.076)	2.027*** (0.071)
High School Complete $\times$ High Tech Imports/GDP	1.513*** (0.080)	1.135*** (0.076)
College Complete $\times$ High Tech Imports/GDP	5.313*** (0.109)	5.420*** (0.105)
GDP Per Capita	0.0003*** (0.000)	0.0003*** (0.000)
Experience	-	0.050*** (0.000)
Experience ²	-0.0006*** (0.000)	-0.0006*** (0.000)
Women	-	-0.372*** (0.001)

Table 3: General Regressions with Fixed Effects, Random Effects, and Year Trend Controls, 1986-2010

Dependent Variable: Log December Real Wage (2005 R\$)	(1) Fixed Effects	(2) Random Effects
North	−0.061*** (0.005)	−0.119*** (0.003)
Northeast	−0.324*** (0.004)	−0.325*** (0.002)
South	−0.158*** (0.004)	−0.041*** (0.002)
Southeast	−0.026*** (0.003)	0.053*** (0.002)
Year Trend	0.021*** (0.000)	−0.030*** (0.000)
Observations	18,350,847	18,350,847
R^2	0.234	0.356
F-statistic	102852.18	
Wald χ^2		2.09×10^6

Note: Standard errors in parentheses. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Men in center west who completed their 8th grade are treated as the base level.

First, I apply empirical models (equation (3) and equation (4)) specified in the empirical framework to the whole dataset to see how technology improvements are correlated with different education levels generally. The random-effects regression suffers from bias due to omitted individual traits. However, since fixed-effects regression looks only at deviations of a variable over time, it cannot give estimates for any time invariant variables. I therefore present results from both fixed-effects regression and random-effects regression. In the fixed-effects regression, gender and education level do not have estimates because they are time invariant. Experience is also dropped because for each individual, his/her work experience change strictly follows the year trend, so experience has been captured by the time index. In the random-effects regression, the estimate of year trend is negative, which could reflect level trend in technological impact on workers who completed 8th grade.

The coefficients of high-tech imports in both regressions are positive. This makes sense because intuitively technology improvements can cause increase in productivity and thus increase in wage. For each regression, combining the coefficient of high-tech imports as share of total GDP with the coefficients of the interaction terms can give us the correlations between high-tech imports as share of total GDP and different education levels' wage returns. Both regressions show that workers who completed 8th grade (medium skill) in the Brazilian formal labor market receive the least benefits from technology improvements comparing with workers with other education levels. Least educated workers and most educated workers benefit more in their wage returns as technology improved. These suggest that as technology improves, the relative wage return of medium-level educations to low-level educations decreases, the relative wage return of high-level educations to medium-level educations increases, which corresponds to Acemoglu and Autor's prediction.

What worth noticing is that both regressions suggest workers who did not complete 4th grade benefits the most under technology improvements, instead of workers who completed college. As a result, the relative wage return of high-level educations to low-level educations decreases under technology improvements, which contradicts Acemoglu and Autor's prediction.

Previous data section suggests that the contradiction may be caused by the omitted consideration of Lula's presidency, given his economic reforms had significant impact on Brazil's general economic conditions, especially on reducing income inequality. Hence, I add a dummy variable indicating Lula's presidency to both regressions and the results are shown in Table A4 in the appendix. After including the presidency of Lula, the results still suggest workers who completed 8th grade (medium skill) in the Brazilian formal labor market receive the least benefits from technology improvements comparing with workers with other education levels. Right now in both regressions, workers who completed college receive the most benefits from technology improvements, instead of workers who did not complete 4th

grade. So, in this case, technology improvements suggest an increase in wage inequality between workers with high-level education and workers with low-level education, just as Acemoglu and Autor predicted. Also, according to Table A4, Lula's presidency benefited people who did not complete 4th grade the most, which further confirmed the unexpected changes in wage inequality between workers with high-level education and workers with low-level education observed in Table 3 can be caused by omitted variable bias.

The unexpected decrease in wage gap between workers who did not complete 4th grade and workers who completed college observed in Table 3 can also be an equilibrium outcome based on market demand and supply in Brazil, which is different from the market in U.S. Table 1 in Section IV shows that workers' education levels in the formal labor market generally increased from 1986 to 2010. Hence, if the demand of low skill jobs remains stable (as shown in Table A3), and the decrease in supply of workers with low education is greater than the increase of workers with medium education shifting to low skill jobs, the wage of workers with low education should increase. If the increase in demand of high skill jobs is less than the increase in supply of highly educated workers, the wage of workers with high education may still decrease even with the skill premium from technology.

VI. Discussions

A. The Role of Trade

Acemoglu and Autor observed the job polarization in the U.S. labor market in the 1970s and I observed similar changes in Brazil started to happen in the 1990s. Given Brazil's trade liberalization ends at the mid 1990s and U.S. is one of the main trade partners of Brazil who provides high-tech exports, it is reasonable to suggest that technology improvements in the past decades first caused job polarization in developed countries and this trend spread to developing countries through multiple channels of technology diffusion, especially through international trade in capital goods embodying high-tech. In my study, I only use this channel as a proxy to technological changes, so I cannot say anything about the significance of its contributions to overall technological changes in Brazil. Future studies can explore other channels such as worker migrations and factory reallocations together with international trade of capital goods to see which channel plays a more important role in technology diffusion and improvements.

B. Future Plans

The main limitation of my current study is that I only used data from the formal labor market in Brazil, which can cause bias in my results given informal labor market is also an important part in Brazil. Moreover, the available variables in my dataset do not include all the things that can affect the wage return to different education levels, such as school quality measures. For future research, I should add informal labor market data to my empirical test and try to add more variables that can be correlated with education return into my model to decrease the bias of my findings.

Moreover, it is reasonable to infer that besides having different impacts on wage returns for workers with different education levels, technology changes may also have different impacts when we change the grouping method. For example, instead of grouping workers by their education levels, we can also group by their genders and create an interaction term between technology and gender in the empirical model to see the difference. Further research about how technology interacts with different properties of workers can help us see the whole picture of the impacts of technology improvements and can help determine winners and losers, which is important for designing future policies.

VII. Conclusion

Inspired by Acemoglu and Autor's study of the U.S. labor market under technology changes, I use this study to check if the trends they observed in U.S. is also observable in Brazil and whether the predictions they made about changes in relative wage among workers with different education levels still hold in the Brazilian labor market.

With values of high-tech imports as share of GDP as the proxy of technology changes, I find technology improvements indeed relatively harms people with medium-level education (8th grade complete) the most in the Brazilian formal labor market. As a result, holding all else constant, the empirical test suggest technology improvements will cause relative return of medium-level education to low-level education to decrease and relative return of high-level education to medium-level education to increase during 1986-2010 in the Brazilian formal labor market, which are consistent with Acemoglu and Autor's predictions.

However, in the Brazil formal labor market, workers who benefited the most from technology improvements are not workers who completed college, but those who did not finish 4th grade. This can be a biased result since after including Lula's presidency into my empirical model, the new coefficients of interaction terms suggest workers who completed college benefit the most from technology improvements, all else remain the same. It is also possible that workers who did not finish 4th grade indeed benefit the most because market demand and supply of labor with different education levels in Brazil is different from U.S., and thus lead to a different equilibrium outcome.

Despite there are still improvements I can do to my datasets and empirical model, my current results suggest that technological changes indeed impact workers with different education levels differently and workers with medium-level skills are the ones who face the most challenges.

References

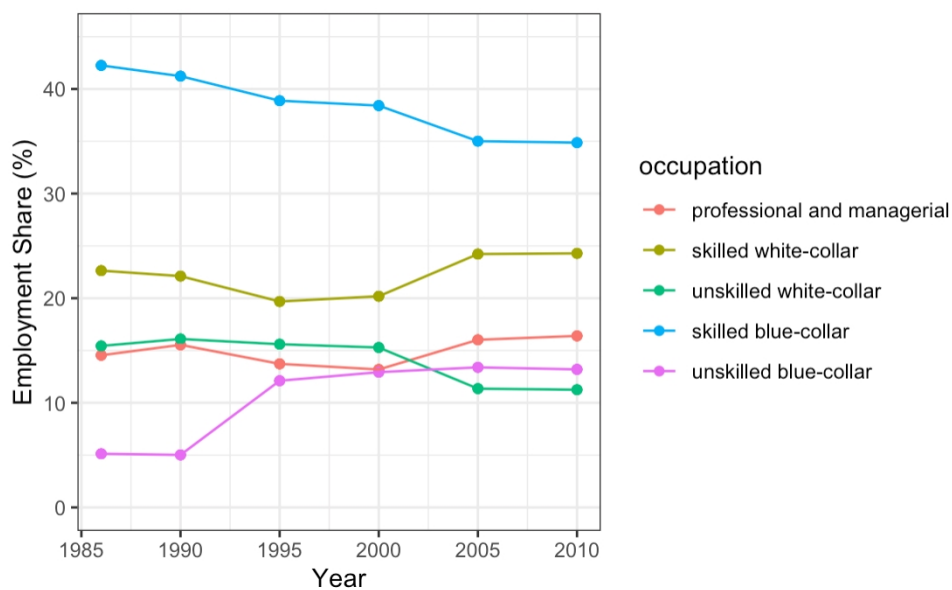
- Acemoglu, D., & Autor, D. (2010). *Skills, Tasks and Technologies: Implications for Employment and Earnings* (Working Paper No. 16082). National Bureau of Economic Research.
<https://doi.org/10.3386/w16082>
- Acemoglu, D., & Autor, D. (2012). *What Does Human Capital Do? A Review of Goldin and Katz's The Race between Education and Technology* (Working Paper No. 17820). National Bureau of Economic Research. <https://doi.org/10.3386/w17820>
- Berman, E., Bound, J., & Machin, S. (1998). Implications of Skill-Biased Technological Change: International Evidence*. *The Quarterly Journal of Economics*, 113(4), 1245–1279.
<https://doi.org/10.1162/0033553985555892>
- Card, D., & Krueger, A. B. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy*, 100(1), 1–40. <https://doi.org/10.1086/261805>
- Connolly, M. (2003). The dual nature of trade: Measuring its impact on imitation and growth. *Journal of Development Economics*, 72(1), 31–55. [https://doi.org/10.1016/S0304-3878\(03\)00067-1](https://doi.org/10.1016/S0304-3878(03)00067-1)
- Ferreira, F. H. G., Firpo, S. P., & Messina, J. (2018). Ageing Poorly?: Accounting for the Decline in Earnings Inequality in Brazil, 1995-2012. IDB Publications (Working Papers), Article 8220.
<https://ideas.repec.org/p/idb/brikps/8220.html>
- Goldin, C., & Katz, L. F. (2010). *The Race between Education and Technology*. Harvard University Press.
- Harmon, C., Oosterbeek, H., & Walker, I. (2003). The Returns to Education: Microeconomics. *Journal of Economic Surveys*, 17(2), 115–156. <https://doi.org/10.1111/1467-6419.00191>

- Hungerford, T., & Solon, G. (1987). Sheepskin Effects in the Returns to Education. *The Review of Economics and Statistics*, 69(1), 175–177. <https://doi.org/10.2307/1937919>
- Jensen, R. (2010). The (Perceived) Returns to Education and the Demand for Schooling*. *The Quarterly Journal of Economics*, 125(2), 515–548. <https://doi.org/10.1162/qjec.2010.125.2.515>
- Keller, W., & Utar, H. (2016). International Trade and Job Polarization: Evidence at the Worker-Level (Working Paper No. 22315). National Bureau of Economic Research. <https://doi.org/10.3386/w22315>
- Mincer, J. (1974). *Schooling, Experience, and Earnings*. *Human Behavior & Social Institutions No. 2*. National Bureau of Economic Research, Inc.
- Nielsen, H. S., Rosholm, M., Smith, N., & Husted, L. (2004). Qualifications, discrimination, or assimilation? An extended framework for analysing immigrant wage gaps. *Empirical Economics*, 29(4), 855–883. <https://doi.org/10.1007/s00181-004-0221-9>
- Park, K. H. (1996). Educational expansion and educational inequality on income distribution. *Economics of Education Review*, 15(1), 51–58. [https://doi.org/10.1016/0272-7757\(95\)00000-3](https://doi.org/10.1016/0272-7757(95)00000-3)
- Sianesi, B., & Reenen, J. V. (2003). The Returns to Education: Macroeconomics. *Journal of Economic Surveys*, 17(2), 157–200. <https://doi.org/10.1111/1467-6419.00192>
- Silveira-Neto, R., & Azzoni, C. R. (2006). Location and regional income disparity dynamics: The Brazilian case*. *Papers in Regional Science*, 85(4), 599–613. <https://doi.org/10.1111/j.1435-5957.2006.00099.x>
- Stanek, C. (2013). The Educational System of Brazil. 10(1).

Appendix

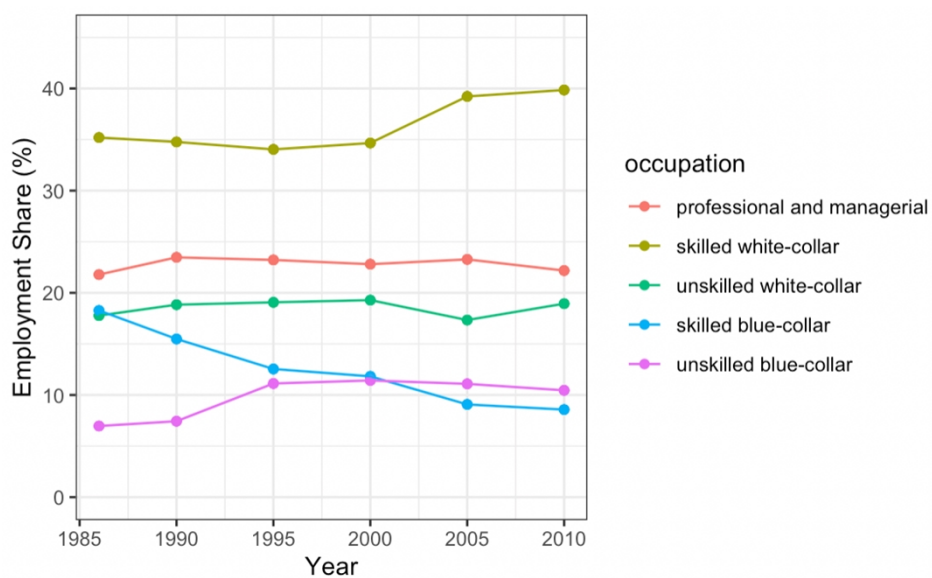
I. Introduction

Figure A1: Employment shares by major occupation groups, Brazil 1986-2010 (Men)



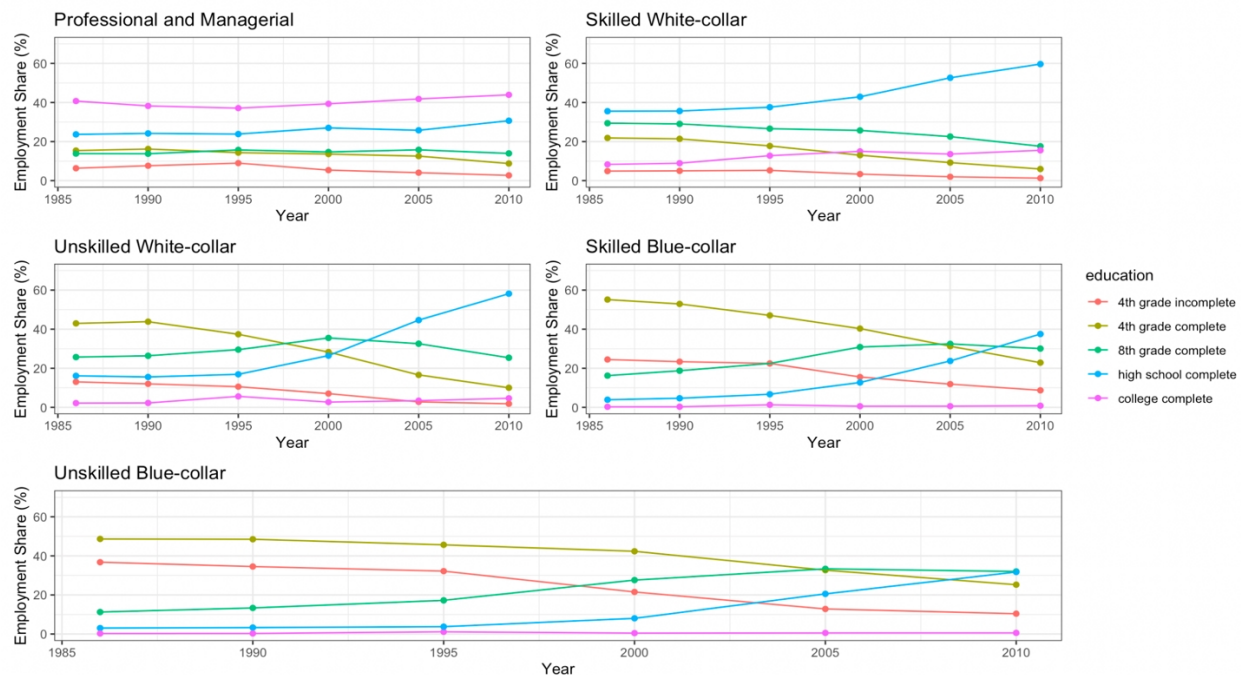
Source: RAIS for years 1986, 1990, 1995, 2000, 2005, and 2010.

Figure A2: Employment shares by major occupation groups, Brazil 1986-2010 (Women)



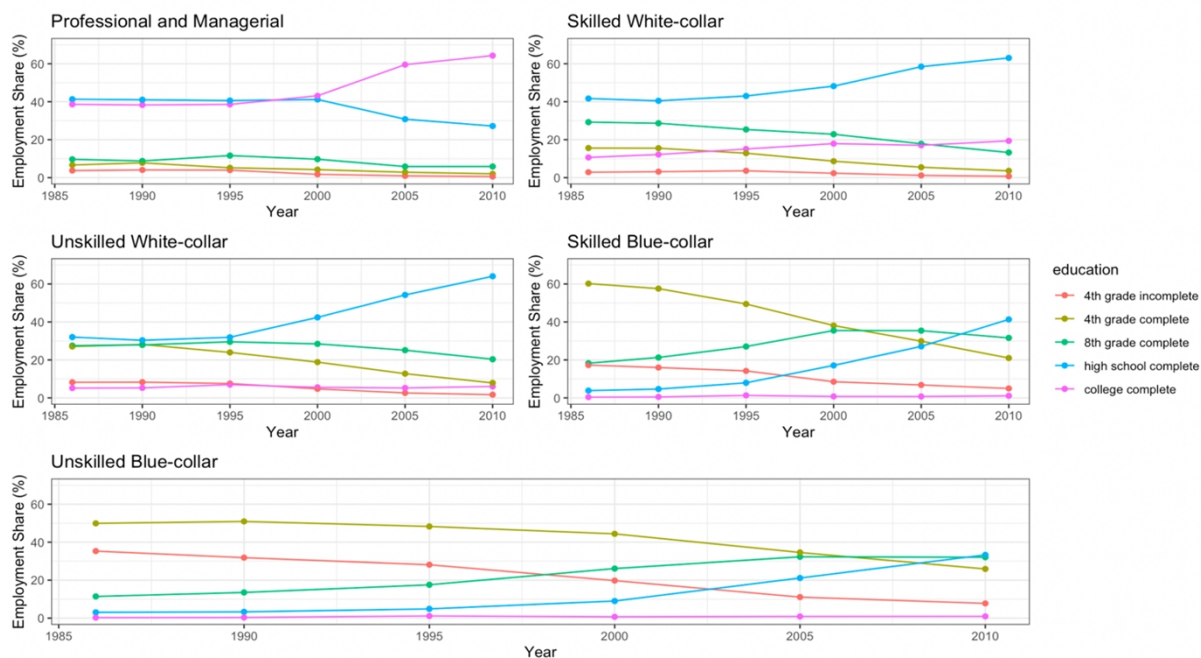
Source: RAIS for years 1986, 1990, 1995, 2000, 2005, and 2010.

Figure A3: Occupational employment shares by worker education, Brazil 1986-2010 (Men)



Source: RAIS for years 1986, 1990, 1995, 2000, 2005, and 2010.

Figure A4: Occupational employment shares by worker education, Brazil 1986-2010 (Women)



Source: RAIS for years 1986, 1990, 1995, 2000, 2005, and 2010.

II. Data

Table A1: Summary Statistics of Real Wage (in 2005R\$) Grouped by Gender

Gender	1986	1990	1995	2000	2005	2010
men	1244.36 (1879.66)	829.22 (1383.37)	1066.08 (1852.17)	1077.28 (1886.37)	1058.23 (1948.65)	1325.21 (2301.03)
women	920.47 (1268.48)	646.36 (1014.70)	867.86 (1431.88)	954.11 (1561.08)	919.61 (1497.59)	1136.49 (1812.22)
Total Observations	650,930	762,303	824,107	878,125	1,086,342	1,232,365

Table A2: Summary Statistics of Real Wage (in 2005R\$) Grouped by Region

Region	1986	1990	1995	2000	2005	2010
Center West	1152.88 (1861.40)	900.74 (1528.31)	1167.95 (2134.87)	1132.91 (2154.82)	1147.57 (2300.96)	1523.48 (2697.56)
North	935.32 (1478.22)	680.87 (1188.58)	991.08 (1740.54)	912.63 (1570.45)	893.99 (1600.18)	1170.05 (1993.86)
Northeast	912.21 (1562.20)	592.17 (1106.89)	721.68 (1405.03)	751.70 (1436.39)	744.67 (1362.06)	985.56 (1710.77)
South	1016.28 (1484.70)	698.92 (1130.13)	908.84 (1570.13)	933.46 (1533.63)	931.03 (1524.32)	1154.36 (1782.60)
Southeast	1247.51 (1802.81)	827.30 (1327.59)	1079.93 (1775.32)	1143.20 (1880.26)	1099.90 (1904.77)	1336.67 (2242.72)
Total Observations	650,930	762,303	824,107	878,125	1,086,342	1,232,365

Resource: Random Sample from RAIS Dataset

Note: numbers inside (.) are standard deviations and numbers above are means.

Table A3: Occupation Distribution (in percentage)

Occupation	1986	1990	1995	2000	2005	2010
Professional and Managerial	15.40	16.15	14.86	15.01	16.88	18.41
Skilled White-collar	25.25	24.59	22.65	23.69	28.32	28.05
Unskilled White-collar	16.46	17.09	16.73	16.68	13.62	13.65
Skilled Blue-collar	36.52	35.64	32.41	30.79	27.62	27.08
Unskilled Blue-collar	6.37	6.53	13.35	13.83	13.56	12.81
Total Observations	561,907	652,490	796,504	854,610	1,076,643	1,224,600

Resource: Random Sample from RAIS Dataset

III. Results

Table A4: General Regressions with Fixed Effects, Random Effects, and Year Trend Controls, 1986-2010, including consideration of Lula's Presidency

Dependent Variable: Log December Real Wage (2005 R\$)	(1) Fixed Effects	(2) Random Effects
Constant	4.389*** (0.004)	3.979*** (0.004)
High Tech Imports/GDP	9.901*** (0.060)	10.02*** (0.058)
4th Grade Incomplete	-	-0.658*** (0.003)
4th Grade Complete	-	-0.319*** (0.002)
High School Complete	-	0.381*** (0.003)
College Complete	-	1.170*** (0.004)
4th Grade Incomplete $\times$ High Tech Imports/GDP	4.426*** (0.096)	4.467*** (0.089)
4th Grade Complete $\times$ High Tech Imports/GDP	1.328*** (0.072)	1.552*** (0.068)
High School Complete $\times$ High Tech Imports/GDP	0.945*** (0.075)	0.707*** (0.072)
College Complete $\times$ High Tech Imports/GDP	4.634*** (0.103)	4.734*** (0.099)
GDP Per Capita	0.0003*** (0.000)	0.0003*** (0.000)
Experience	-	0.051*** (0.000)

Table A4: General Regressions with Fixed Effects, Random Effects, and Year Trend Controls, 1986-2010, including consideration of Lula's Presidency

Dependent Variable: Log December Real Wage (2005 R\$)	(1) Fixed Effects	(2) Random Effects
Experience ²	−0.0006*** (0.000)	−0.0006*** (0.000)
Women	-	−0.372*** (0.001)
Lula's Presidency	−0.048*** (0.001)	−0.044*** (0.001)
4th Grade Incomplete × Lula's Presidency	0.118*** (0.002)	0.126*** (0.002)
4th Grade Complete × Lula's Presidency	0.023*** (0.001)	0.029*** (0.001)
High School Complete × Lula's Presidency	0.029*** (0.001)	0.020*** (0.001)
College Complete × Lula's Presidency	0.034*** (0.002)	0.033*** (0.002)
North	−0.061*** (0.005)	−0.119*** (0.003)
Northeast	−0.323*** (0.004)	−0.325*** (0.002)
South	−0.158*** (0.004)	−0.041*** (0.002)
Southeast	−0.026*** (0.003)	0.054*** (0.002)
Year Trend	0.023*** (0.000)	−0.028*** (0.000)
Observations	18,350,847	18,350,847
R^2	0.234	0.356
F-statistic	82192.67	
Wald χ^2		2.23×10^6

Note: Standard errors in parentheses. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Men in center west who completed their 8th grade
and not in Lula's ruling period are treated as the base level.